

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-071688

(43)Date of publication of application : 12.03.2002

(51)Int. Cl. G01N 33/53
C12N 15/09
C12Q 1/68
G01N 21/64
G01N 21/78
G01N 33/566
G01N 37/00

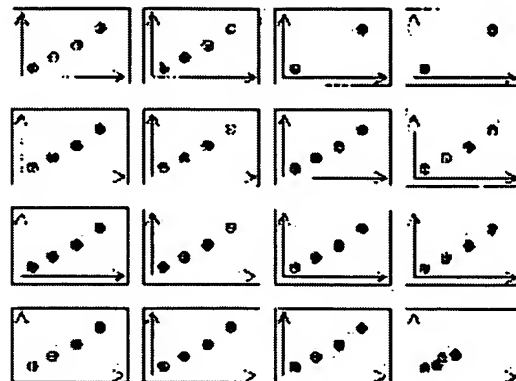
(21)Application number : 2000-265933 (71)Applicant : HITACHI SOFTWARE ENG CO LTD
(22)Date of filing : 01.09.2000 (72)Inventor : WATANABE TSUNEHICO
NOZAKI YASUYUKI
NAKASHIGE AKIRA
TAMURA TAKURO

(54) EXPERIMENT RESULT DISPLAY METHOD AND EXPERIMENT ERROR EVALUATION METHOD
FOR HYBRIDIZATION REACTION USING BIOTIP

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a visual display effective to grasp nonconformity in a manufacturing and experimenting processes from the data of biotip after an experiment.

SOLUTION: The same sort of control spotted on each partition on one biotip is measured and plotted on one graph every partition. All of these are plotted on one screen in conformity to the partition positions of all the biotips. The respective graphs are all displayed on one screen to allow an experiment error as the whole of one biotip, and the experiment error can be determined by focusing on linearity and angle of elevation as to the dispersion degree of control.



LEGAL STATUS

[Date of request for examination] 04.10.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998, 2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2002-71688

(P2002-71688A)

(43) 公開日 平成14年3月12日 (2002.3.12)

(51)Int.Cl. ⁷	識別記号	F I	テマコード*(参考)
G 0 1 N 33/53		G 0 1 N 33/53	M 2 G 0 4 3
C 1 2 N 15/09		C 1 2 Q 1/68	A 2 G 0 5 4
C 1 2 Q 1/68		G 0 1 N 21/64	F 4 B 0 2 4
G 0 1 N 21/64		21/78	C 4 B 0 6 3
21/78		33/566	
審査請求 未請求 請求項の数6 O L (全 12 頁) 最終頁に続く			

審査請求 未請求 請求項の数 6 O L (全 12 頁) 最終頁に続く

(21) 出願番号	特願2000-265933(P2000-265933)	(71) 出願人	000233055 日立ソフトウェアエンジニアリング株式会社 神奈川県横浜市中区尾上町6丁目81番地
(22) 出願日	平成12年9月1日(2000.9.1)	(72) 発明者	渡辺 恒彦 神奈川県横浜市中区尾上町6丁目81番地 日立ソフトウェアエンジニアリング株式会社内
		(74) 代理人	100091096 弁理士 平木 祐輔 (外1名)

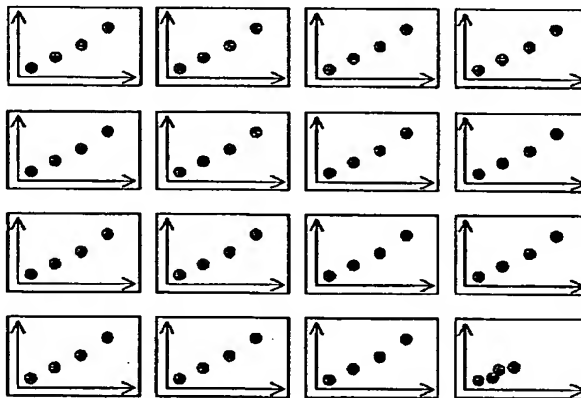
最終頁に続く

(54) 【発明の名称】 バイオチップを用いたハイブリダイゼーション反応の実験結果表示方法及び実験誤差評価方法

(57) 【要約】

【課題】 実験後のバイオチップのデータから作製・実験工程における不具合を把握するために有効なビジュアル表示を提供する。

【解決手段】 一枚のバイオチップ上の各区画にスポットしてある同一種類のコントロールを計測し、それを区画ごとに一つのグラフ上にプロットし、これらをすべてバイオチップにおける区画の位置に合わせて一つの画面上にプロットする。それぞれのグラフを一つの画面上にすべて表示することにより、一枚のバイオチップ全体としての実験誤差の把握が可能となる。また、コントロールのばらつき度合いに関し、直線性と仰角に着目することにより実験誤差の定量化が可能となる。



【特許請求の範囲】

【請求項1】 複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに前記2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットし、各区画ごとのグラフを前記バイオチップ上に設定されたそれぞれの区画の位置に合わせて一つの画面上に表示することを特徴とするバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法。

【請求項2】 複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに前記2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットの近似直線と各プロットとの決定係数を求め、各区画の決定係数を区画に対応させてグラフ表示することを特徴とするバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法。

【請求項3】 複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光色素の蛍光シグナル計測データを各区画ごとに前記2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットと原点とを結んだ直線と横軸との仰角の最大値、最小値、平均値を各区画に対応させてグラフ表示することを特徴とするバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法。

【請求項4】 複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに前記2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットの近似直線と各プロットとの決定係数を用いて実験誤差を評価することを特徴とするバイオチップを用いたハイブリダイゼーション反

応の実験誤差評価方法。

【請求項5】 複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに前記2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットと原点とを結んだ直線と横軸との仰角を用いて実験誤差を評価することを特徴とするバイオチップを用いたハイブリダイゼーション反応の実験誤差評価方法。

【請求項6】 請求項5記載のバイオチップを用いたハイブリダイゼーション反応の実験誤差評価方法において、

仰角の最大値と最小値と平均値を用いることを特徴とするバイオチップを用いたハイブリダイゼーション反応の実験誤差評価方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、特定の遺伝子とハイブリダイズさせることによって得られた遺伝子発現データの表示及び評価に関し、そのデータを得る実験過程における実験の失敗・誤差を検出するときに視覚的にわかりやすい形式によって表示するための表示方法及び評価方法に関する。

【0002】

【従来の技術】ゲノム配列が決定された種の増加に伴い、進化に対応すると見られる遺伝子を見つけ出し、どの生物も共通に持っていると考えられる遺伝子の集合を探索する、また、逆に種に個別な特徴を推測するため、種間の遺伝子の違いから何かを見出そうとする、いわゆるゲノム比較法が盛んに行われてきた。

【0003】しかし近年、バイオチップやDNAチップ（以下、バイオチップという）に代表されるインフラストラクチャの発達によって、分子生物学の興味は、種間の情報から種内の情報へ、すなわち同時発現解析へと移りつつあり、これまでの種間の比較と合わせて、情報の抽出から関連付けの場が大きく広がりを持ち始めている。

【0004】例えば、既知の遺伝子と同一の発現パターンを示す未知の遺伝子が見つければ、その遺伝子には既知の遺伝子と同様の機能があると類推できる。これら遺伝子や蛋白質そのものの機能的な意味付けは、機能ユニットや機能グループといった形で研究されている。またそれらの間の相互作用も、既知の酵素反応データや物質代謝データとの対応付けによって、あるいはより直接的に、ある遺伝子を破壊あるいは過剰反応させ、その遺伝子の発現をなくすか、あるいは多量に発現させ、その遺

伝子の直接的及び間接的影響を、全遺伝子の発現パターンを調べることによって解析している。

【0005】遺伝子の発現パターンを調べる時、バイオチップを用いた実験では、まず調べたい生体組織に関するエレメントを用意しておく。ここで、エレメントとは実験対象の生体組織に関係するDNAを断片化したものであり、このエレメントを、スライドガラスやシリコンなどの基板上に、1平方センチメートル当たり数百から数千の密度でスポッティングし固定化したものがバイオチップである。また、サンプルとは実験対象の生体組織から抽出したDNAを断片化したもの又はRNAであり、バイオチップ上のエレメントと反応させるものを指す。細胞中の遺伝子が発現するときDNAはRNAに転写される。このRNAを抽出し、蛍光標識してサンプルとする。サンプルとエレメントとを反応させると、相補的な関係にある一本鎖同士が互いに結合されハイブリダイゼーションする。バイオチップでは、ハイブリダイゼーションを用い、生体組織での遺伝子の発現状態を定量的又は定性的に検出することができる。

【0006】この分野において成功した事例として、東大医科研による薬の有効性に関する実験結果の報告がある(T. Tsunoda et al.: Discrimination of Drug Sensitivity of Cancer Using cDNA Microarray and Multivariate Statistical Analysis: Genome Informatics 1999 (1999, Dec.) pp.227-228, Universal Academy Press Inc.)。実験では、正常細胞から抽出したRNAと癌細胞から抽出したRNAを各々異なる色の蛍光色素で標識した上で混合し、バイオチップ上でエレメント(遺伝子)とハイブリダイズさせた後、双方の蛍光色素から発生される蛍光シグナルの強度を測定している。

【0007】図16は、この実験によって得られる各遺伝子の発現状態の表示方法を模式的に示した図である。この表示はバイオチップ上の遺伝子にハイブリダイズした蛍光シグナルのデータをプロットしたもので、片方の軸に正常細胞、他方の軸に癌細胞の蛍光シグナルをとっている。図中の一つのプロットが一つの遺伝子に対応する。データの分析においては、ある一定以上の蛍光シグナルが出ている遺伝子に関して、正常細胞のシグナルと癌細胞のシグナルの比に着目して、疾病に特異的な遺伝子の候補を絞り込んでいる。具体的には、図16の領域Aに入る遺伝子(正常細胞中で働いているが癌細胞中では働かなくなるもの)と領域Bに入る遺伝子(癌細胞中で働くが正常細胞中では働かなくなるもの)を特に選出している。このような表示方法をとることで、特定の疾病に特異的に働く遺伝子の候補を絞り込むことが可能となっている。

【0008】

【発明が解決しようとする課題】上記のようなバイオチップのデータ分析を行うためには、そのデータ自体の信頼性が保証されたものである必要がある。すなわち同一

環境のもとに実験を行ったとき、結果に再現性がなければならない。しかし、実際のバイオチップ作製とバイオチップを用いた実験の一連の技術は、発展途上の段階で再現性が完全には保証されていない状況である。この原因としては、バイオチップにスポットするエレメントの分量が一定でないこと、温度や湿度などの環境の変化に技術が対応しきれていないことなどが挙げられる。さらに、バイオチップ上でハイブリダイゼーションの反応効率を一定に維持する技術や、ハイブリダイゼーション後の蛍光の読み取り精度等に関しても実験技術的に完全に確立されているとは言い難い。つまり、これら一連の実験より得られるデータにどこまで信頼性があるのか不明な点が多いのが現状である。

【0009】図17は、バイオチップの実験結果をスキャナで読み取った時のイメージデータの模式図である。今まで研究者は、このようなバイオチップの読み取りイメージを目視して、そのバイオチップによる実験データの採用・不採用を判断していた。例えば、読み取ったイメージデータが全体的に暗いとき(発現が観測されなかったとき)や一部のみ明るいとき、一部しか発現しなかったときなどに、そのバイオチップの実験データを不採用と判断していた。このようなことは、ハイブリダイゼーション反応が成功していなかったときや、バイオチップの基板に傷が入っていたとき、バイオチップ上のスポットの量が場所によって均一でなかったりしたときに起こるとされているが、明確な原因は突き止められていない。

【0010】現在、バイオチップを作製する側からすると、バイオチップ作製工程における技術の精度を向上させ、誤差を抑えて信頼性の高いバイオチップを大量に作製できる技術を確立することが急務である。そのためには、バイオチップの作製精度あるいは作製誤差を正確に知るための評価方法あるいはツールが必要である。一方、バイオチップを使用して実験を行う研究者の側からすれば、バイオチップの実験結果を検証することにより、その実験結果の採用・不採用を決定し、不採用の場合にはその原因の詳細を突き止めるための評価方法あるいはツールがあれば便利である。このように、バイオチップの作製・実験工程におけるどの部分で、どのような不具合が生じたのかを知り、結果をその後のバイオチップ作製あるいは実験に反映させるための評価方法が求められている。本発明は、このようなバイオチップ製造者及びバイオチップ利用者双方の要望に応え、実験後のバイオチップのデータからバイオチップの作製工程あるいは実験工程における不具合を把握するために有効な方法を提供することを目的とする。

【0011】

【課題を解決するための手段】本発明では、バイオチップの読み取りデータに含まれる誤差をビジュアル表示すること、また、その誤差を定量化することによって前記

目的を達成する。具体的には、1枚のバイオチップ上に設定された複数の区画のそれぞれに希釈度の異なる同一種類のコントロールを複数スポットし、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルを用いてハイブリダイゼーション反応を行う。そして、ハイブリダイゼーション反応後のコントロールから得られる2種類の蛍光シグナル計測データを区画ごとに一つのグラフ上にプロットし、これらをバイオチップ上に設定されたそれぞれの区画の位置に合わせて一つの画面上に表示して比較する。また、コントロール計測値のばらつき度合いに関し、縦軸と横軸にそれぞれの蛍光シグナル強度をとって各コントロールの計測値をプロットしたときのプロット配列の直線性、あるいはプロット配列を近似する直線の角度を調べることで、実験誤差を定量化する。

【0012】バイオチップを用いた実験では、バイオチップ上のスポットの量や、スポットに含まれるDNA、RNA、cDNA等のエレメントの量、ハイブリダイゼーションの反応の差などの理由から、バイオチップ間あるいはバイオチップ内の位置によって、観測される蛍光シグナルと本来の発現の度合いとの間にずれが生じることがある。そこで、それらの差を補正するために、コントロールと呼ばれるものをバイオチップ上に配置する。コントロールとしては、例えば、個体を維持するために各細胞で一定の発現をしているハウスキーピングとよばれる遺伝子が用いられる。この他にも、動物の実験に対して植物だけが発現しないような必ず発現しないことが保証されている遺伝子や、遺伝子とは関係のない蛍光色素などが用いられる。これをスポットとしてバイオチップに載せて蛍光シグナルの基準値とする。通常、コントロールは蛍光シグナルの基準値としてデータを補正するために用いられるが、本発明ではデータのばらつき度合いを測るために用いる。

【0013】本発明では、バイオチップ自体の実験誤差の検出にコントロールの計測データを使用し、それらを区画ごとに一つのグラフ上にプロットし、これらをすべてバイオチップの位置に合わせて一つの画面上に並べて表示する方式をとる。また、本発明ではこのコントロールの計測データのばらつき度合いを定量化するため、二つの方法を用意する。一つはコントロール計測データの直線性に着目したものである。すなわち、コントロールの希釈度が異なっても2種類の蛍光シグナルの強度比が一定になるという仮定のもと、異なる希釈度を有する複数のコントロール計測データのプロットに近似する直線を求め、決定係数とよばれるプロットした点が直線に近いかどうかを表す尺度で定量的に考察する。近似直線の決定係数を測定することにより評価を行い、誤差を定量化することが可能になる。もう一つは、コントロール計測データのプロットのグラフ上での仰角に着目したものである。すなわち、原点からプロットした点の仰角を

求めることにより誤差を定量化する。

【0014】これらの結果から、バイオチップの作製・実験工程における実験条件の変化などと照らし合わせることで、バイオチップの実験工程のどこでミスが生じたのか推測することが可能となる。例えば誤差の原因となるものとして、気温、湿度等の環境によるスポットの液量の変化、ハイブリダイゼーション反応の不均一性、ハイブリダイゼーション反応後のバイオチップの洗浄不足、スポットからの蛍光を計測するとバイオチップ基板の傾きによる蛍光測定装置のスキャニング誤差、バイオチップ基板の歪み、空気中・溶液中に存在する塵によるスキャニングの誤差、バイオチップ基板がもともと発する蛍光、スキャナの光電子増倍管のノイズ等が挙げられる。これらと、本発明によって得られる誤差の定量化した値とを関連付けることにより、再度実験を行ったとき、その実験結果から誤差の原因が何であるかをきき止めるための推測を支援することが出来る。

【0015】すなわち、本発明によるバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法は、複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットし、各区画ごとのグラフをバイオチップ上に設定されたそれぞれの区画の位置に合わせて一つの画面上に表示することを特徴とする。

【0016】本発明によるバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法は、また、複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区画ごとに2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットの近似直線と各プロットとの決定係数を求め、各区画の決定係数を区画に対応させてグラフ表示することを特徴とする。

【0017】本発明によるバイオチップを用いたハイブリダイゼーション反応の実験結果表示方法は、また、複数の区画に分割されたスポット領域のそれぞれの区画に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールか

ら得られる2種類の蛍光色素の蛍光シグナル計測データを各区分ごとに2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットと原点とを結んだ直線と横軸との仰角の最大値、最小値、平均値を各区分に対応させてグラフ表示することを特徴とする。

【0018】本発明によるバイオチップを用いたハイブリダイゼーション反応の実験誤差評価方法は、複数の区分に分割されたスポット領域のそれぞれの区分に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区分ごとに2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットの近似直線と各プロットとの決定係数を用いて実験誤差を評価することを特徴とする。

【0019】本発明によるバイオチップを用いたハイブリダイゼーション反応の実験誤差評価方法は、また、複数の区分に分割されたスポット領域のそれぞれの区分に同一種類のコントロールが希釈度を変えて複数スポットされているバイオチップを用い、各々異なる蛍光色素で標識した2種類のサンプルを混合した混合サンプルによるハイブリダイゼーション反応後に前記コントロールから得られる2種類の蛍光シグナル計測データを各区分ごとに2種類の蛍光色素の蛍光シグナル強度を縦軸と横軸にとった一つのグラフ上にプロットしたとき得られる当該プロットと原点とを結んだ直線と横軸との仰角を用いて実験誤差を評価することを特徴とする。仰角としては、その最大値と最小値と平均値を用いることが好ましい。

【0020】

【発明の実施の形態】以下、図面を参照して本発明の実施の形態を説明する。図1は、本発明のシステム構成例を示す図である。本発明のシステムは、一連の細胞のプロセスにおいて遺伝子の発現の度合いを数値化したものを遺伝子発現データとして格納している記憶装置100、その発現データを視覚化して表示するための表示装置101、本システムへの値の入力や選択の操作を行うためのキーボード102やマウス103などの入力装置、コントロールのデータの値より実験誤差の定量化処理を行う処理部104から構成される。処理部104では、グラフ上におけるコントロールスポットのプロット、誤差の定量化処理（直線性、仰角の計算）を行う。

【0021】図2に、記憶装置100に格納されている遺伝子発現データの具体例を示す。このデータは、各種の遺伝子に対して正常の細胞Aと疾病に罹患した細胞Bとを比較した実験データを示している。実験の結果は、遺伝子IDをインデックスとする発現量（各細胞に標識した

蛍光シグナルの測定値）として表にまとめてある。表中の数値は、例えば遺伝子ID1番の遺伝子に関して、バイオチップ上のハイブリダイゼーション反応で正常細胞Aの蛍光シグナルが1,234という数値で計測され、疾病に罹った細胞Bの蛍光シグナルが56という数値で計測された、ということを表している。対象遺伝子の総数は実験により異なるが、現在のバイオチップでは数百のものから数万のオーダーのものまである。

【0022】図3は、バイオチップの一例を示す模式図である。一枚のバイオチップ300はいくつかの区分301に区切られている。図示したバイオチップ300の例では、縦横4つずつ区切られた合計16の区分に分割されている。そして、それぞれの区分301には、複数のコントロールのスポット302と、サンプルとハイブリダイズする遺伝子、DNA断片、RNA等のエレメントのスポット303が用意されている。一枚のバイオチップ300上の全てのコントロールのスポット302には同じコントロール物質がスポットされている。コントロールとしては、前述のようにハウスキーピング遺伝子、必ず発現しないことが保証されている遺伝子、蛍光色素などを用いることができる。

【0023】コントロールは何種類かの濃度に希釈してスポットする。図4には、コントロールを4種類の濃度（コントロールの原液、10分の1希釈液、100分の1希釈液、1000分の1希釈液）に分けて、それらに対応する蛍光シグナルのデータをグラフにプロットしたものを示している。このときコントロールには、正常細胞か否かに関わらず一定の発現量を示すことが知られている既知の遺伝子などを利用するので、蛍光シグナルデータは本来ならば図4に示すような45度の直線上に、希釈の度合いに応じた間隔で並ぶべきである。なぜなら、図4では、Y軸に正常細胞のサンプルに標識した蛍光色素のシグナルをとり、X軸に癌細胞のサンプルに標識した蛍光色素のシグナルをとっているが、コントロールはどちらの細胞においても均等に含まれる遺伝子を用いるため2種類の蛍光シグナルの強度は均等になるからである。

【0024】ハイブリダイゼーション実験は、図3に示したバイオチップを用い、2種類の細胞から抽出したそれぞれ異なる蛍光色素で標識したRNA、例えば正常細胞から抽出したRNAとガン細胞から抽出したRNAを等量ずつ混合したものをサンプルとして用いて行う。ハイブリダイゼーション終了後に、バイオチップに励起光を照射し、バイオチップの各区分にプロットされているコントロールスポット及びエレメントスポットから放出される2種類の蛍光シグナル強度を測定し、遺伝子発現データとして取り込む。

【0025】図5は、本発明による遺伝子実験データ表示方法の一例の概略処理フローを示した図である。この処理フローに従って順次説明する。まず、ステップ500

において、図1に示す記憶手段100から処理部104へ遺伝子発現データを読み込む。次に、ステップ501において、バイオチップ上のコントロールデータを区画ごとにプロットし、それを一つの画面上にバイオチップ上の各区画に対応した形で表示する。

【0026】例えば図3に示したバイオチップでは、縦横4つずつの合計16の区画に分割されており、各区画にはそれぞれ複数のコントロール302がプロットされている。各区画にプロットされているコントロールは全て同一種類のコントロールである。ここでは、各区画を一意に決めるため、区画IDとして図6に示すように左からa番目、上からb番目の区画を(a, b)区画と定義することにする。各区画毎に濃度の異なる各コントロールから得られる2種類の蛍光シグナルを、一方の軸に正常細胞から抽出したRNAに標識した蛍光色素からの蛍光シグナル強度を取り、他方の軸に癌細胞から抽出したRNAに標識した蛍光色素からの蛍光シグナル強度をとって、プロットしグラフ表示する。

【0027】このとき、図7あるいは図8に示すように、一つの画面上に一枚のバイオチップの各区画のグラフをバイオチップ上の区画と同様の順番に表示する。この表示により、バイオチップ上のどの区画においてどのような反応がおこっているかを視覚的に確認し、一枚のバイオチップ全体としての実験誤差の発生状況を把握するのに有効なビジュアル表示を提供できる。例えば、図7に示すように、全ての区画について、プロットが傾き約45°の直線に乗る同様なグラフが得られたとすれば、バイオチップの製造に問題はなく、また、ハイブリダイゼーション反応は各区画で均一に生じていると推定できる。一方、図8に示すように、特定の区画(図の場合、右下の区画)のコントロールに対する計測結果が他の区画と異なった傾向を示している場合には、バイオチップ上の不具合が右下の区画(4, 4)において生じていることを検出できる。

【0028】図5に戻り、これらの表示に対して誤差の定量化を行う場合(図5のステップ502)には、定量化の方法を選択する(図5のステップ503)。直線性に着目した方法で定量化する場合、ステップ504, 505の処理を行う。まず、コントロールの希釈濃度によらず2種類の蛍光シグナルの比が一定になるという仮定のもと、異なる希釈率の複数のコントロールのプロットに近似する直線を最小二乗法により求める。そして、複数のプロットが直線に近いかどうかを表す決定係数とよばれる尺度を用い、直線性を定量的に考察する。ここで、最小二乗法とは、グラフの点列の分布を直線・曲線・平面などの式によって近似する方式である。

【0029】図9に示すように、前記近似直線を用いて蛍光色素Aのデータから蛍光色素Bを予測することを考える。すなわち、実測値から縦軸方向に降ろした点を予測値とする。このとき、点の総数をn、実測値の座標を

(x_i, y_i) ($i=1, 2, \dots, n$)、予測値の座標を次の【数1】とし、 y_i ($i=1, 2, \dots, n$)の平均値を【数2】とすると、関係式【数3】が成立する。

【0030】

【数1】

$$(x_i, \bar{y}_i) (i=1, 2, \dots, n)$$

【0031】

【数2】

\bar{y} (全体の平均値)

【0032】

【数3】

$$\sum_{i=1}^n (y_i - \bar{y})^2 = \sum_{i=1}^n (\bar{y}_i - \bar{y})^2 + \sum_{i=1}^n (y_i - \bar{y}_i)^2$$

【0033】これは、実測値 (x_i, y_i) と全体の平均値との誤差は、予測値【数4】と全体の平均値との誤差と、観測値と予測値との誤差の和になることを示している。

【0034】

【数4】

$$(x_i, \bar{y}_i)$$

【0035】そこで、通常あてはまりの良さを評価する尺度として、決定係数とよばれる量が用いられる。決定係数 R^2 とは以下の式で定義される値である。 R^2 は0から1までの範囲をとる値で、 R^2 が1に近いほど当てはまりがよいと判断することができる。

【0036】

【数5】

$$R^2 = \frac{\sum_{i=1}^n (\bar{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2}$$

【0037】また、蛍光シグナルBのデータから蛍光シグナルAを予測する近似直線でも、上と同様のことが成立する。このとき定義される決定係数は、蛍光シグナルAのデータから蛍光シグナルBを予測する近似直線の決定係数 R^2 と等しいことが知られている。ここでは最小二乗法を用いて直線を近似する例を説明したが、この他にも図10に示すように各点から直線に下ろした垂線の長さの和が最小になるように直線を決定する方法もある。この時も最小二乗法と同様に決定係数を定義することができる。

【0038】図11に、直線性を用いて定量化した誤差の結果表示の具体例を示す。この図は、縦軸に決定係数を、横軸に区画IDをとり、区画間のコントロールの傾向を調べたものである。図11(a)に示すコントロール

Aの例では、決定係数はどの区画においてもほぼ1に近い値を示しており、これらのコントロールでは異なる希釈のものを並べたときにグラフ上のプロットは直線状になることを示している。これに対し、図11(b)に示したコントロールBの例では、区画間で決定係数にばらつきがあり、蛍光シグナルAと蛍光シグナルBの比の値が安定しないことがわかる。

【0039】一方、仰角に着目した方法で定量化する場合、図5のステップ506, 507の処理を行う。希釈度の異なるコントロールについてのデータをグラフ上にプロットしたとき、同じコントロール物質であれば蛍光シグナルAと蛍光シグナルBの比が一定になり、各コントロールのプロットは図4に示すように45度の直線上に並ぶはずである。そこで図12に示すように、バイオチップの各区画毎に、異なる希釈度を有する複数のコントロール計測データのプロット（図にはコントロール数が4つの場合を示す）の原点に対する最小の仰角・最大の仰角及び仰角の平均値を求める。

【0040】図13に、仰角を用いた定量化の結果表示の具体例を示す。図13のグラフは縦軸に仰角の角度をとり、横軸に区画IDをとり、区画間のコントロールの傾向を表示したものである。図をみるとわかるように、図13(a)に示したコントロールCの例では、どの区画でも仰角の最大と最小の幅が大きい。これは、コントロールから発せられる2種類の蛍光シグナルの比が一定していないことを表している。また、図13(b)に示すコントロールDの例では、どの区画でも最大と最小の幅は小さいのでコントロールから発せられる2種類の蛍光シグナルの比は一定している。しかし、左側の区画ほど仰角が大きく、また下方の区画ほど仰角が大きい傾向があることがわかる。

【0041】バイオチップの実験者は、コントロール計測データをグラフ上にプロットしたときの直線性や仰角を定量化したこれらの表示をもとに、実験段階のどこに誤差が生じているのか推測することができる。例えば、図13(b)のような結果表示が得られたときは、バイオチップがスキャニング時に傾いていたために、左下の区画になるほど一方の蛍光シグナルが本来よりも強く観測され、右上の区画になるほど他方の蛍光シグナルが本来よりも強く観測され、結果として仰角が上下したことが考えられる。つまり、バイオチップ上のスポットは正確にうたれているが、蛍光を測定する位置の物理的な違いにより測定値が変化したことが推測される。

【0042】また、図12のスカッパプロットの例を見ると、コントロール計測データに大きなばらつきがあり、グラフ上でのプロットの直線性は低く、仰角の最大値と最小値との差も大きいことがわかる。これに対し、図14の例ではコントロールのデータはグラフ上で直線上に並んでいるので直線性が高いことがわかる。また、仰角の最大値と最小値の差は小さいので二つの蛍光シグ

ナルの比はほぼ一定している。しかし、直線全体が45度よりも縦軸の方向にずれていることがわかる。このように直線性、仰角の両方に着目することでコントロールの状態の詳細な評価が可能となる。

【0043】以上のように、図3に示すようなバイオチップ上の各区画にスポットしてある同一種類のコントロールの測定結果を、図7あるいは図8に示すようにバイオチップ上の区画に合わせて一つの画面上にすべて表示することで、バイオチップ全体としての実験誤差の把握が可能となる。また、コントロールのばらつき度合いに関し、図9あるいは図10に示すようにグラフ上に配列されたスポットの直線性、図12あるいは図14に示すようにグラフ上に配列されたスポットの仰角を調べることにより、実験誤差を定量化することが可能となる。

【0044】この処理を行うにあたって図15に示すようなインタフェースを用いると操作が容易になる。図15に示すインタフェースでは、表示装置に表示されるウィンドウ1500上に3つの処理をスムーズに行うためのボタンを用意した。ボタン1501は全ての区画のコントロールのデータをプロットすることに対応し、ボタン1502は誤差の定量化において直線性の計算を行うことに対応し、ボタン1503は誤差の定量化における仰角の計算に対応する。

【0045】はじめにボタン1501をマウスなどのポインティングデバイスを用いてクリックする。すると、ウィンドウ1500上の表示枠1505には図7に示すように一枚のバイオチップ上の各区画に対応したそれぞれのスカッパプロットが表示される。次に、誤差の定量化ボタンに関して、ボタン1502をクリックすると上述した直線性の計算を行い、表示枠1505内のそれぞれの区画ごとに近似直線が表示される。これに加え、図11に示したような縦軸に決定係数、横軸に区画IDをとったグラフがウィンドウ1506に表示される。また、ボタン1503をクリックすると、上述したように仰角の計算を行い、図13に示した縦軸に仰角の角度、横軸に区画IDをとったグラフがウィンドウ1507に表示される。

【0046】なお、本発明による実験結果表示方法あるいは実験誤差評価方法は、コンピュータによって実行することができる。すなわち、これらの処理を行うためのプログラムを記録媒体に記録しておき、コンピュータに記録媒体から当該プログラムを読み取らせることによって実行することができる。

【0047】

【発明の効果】本発明によると、バイオチップ上の実験データを視覚的にわかりやすく、バイオチップの実験工程のどこでミスが生じているのか推測しやすい形式によって表示することができる。また、ここで得られるグラフを直線性と仰角に着目して分析することによって実験誤差の定量化が可能となる。

【図面の簡単な説明】

【図1】本発明のシステム構成例を示す図。
 【図2】遺伝子発現データ的具体例を示す図。
 【図3】バイオチップのスポット例を示す模式図。
 【図4】標準的コントロールデータの表示例を示す図。
 【図5】概略処理フローを示す図。
 【図6】バイオチップの区画表示例を示す図
 【図7】一枚のバイオチップにおけるコントロールのデータの表示例を示す図。
 【図8】一枚のバイオチップにおけるコントロールのデータの表示例を示す図。
 【図9】コントロールデータの直線性に関する誤差の定量化の方法を説明する図。
 【図10】コントロールデータの直線性に関する誤差の定量化の他の例を説明する図。
 【図11】コントロールデータの直線性に関する誤差の定量化の結果表示例を示す図。
 【図12】コントロールデータの仰角に関する定量化処理の説明図。
 【図13】コントロールデータの仰角に関する誤差の定量化の結果表示例の図。

【図14】コントロールデータの直線全体に対する仰角に関する定量化の表示例を示す図。

【図15】一枚のチップにおけるコントロールのデータの表示とコントロールデータの直線性に関する誤差の定量化と仰角に関する誤差の定量化の結果表示の為のインタフェース例の図。

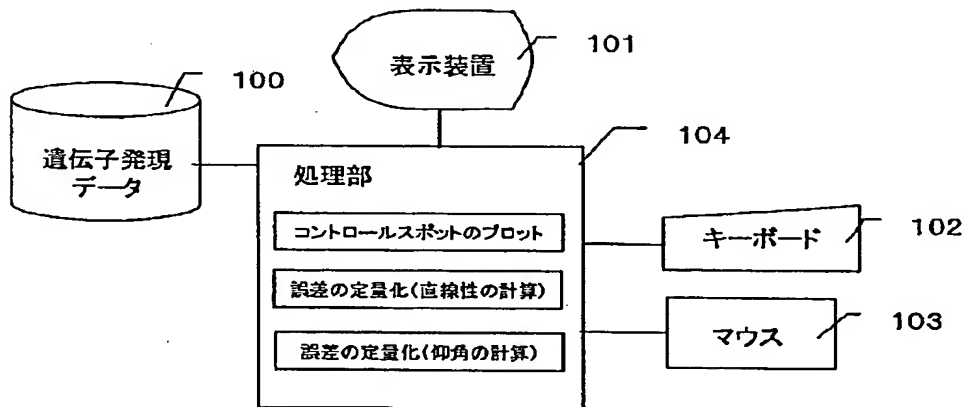
【図16】標準的遺伝子発現データ分析結果表示例の図。

【図17】スキャナから読み取ったバイオチップのイメージ例を示す図

【符号の説明】

100 遺伝子発現データ記憶装置
 101 表示装置
 102 キーボード
 103 マウス
 104 処理部
 300 バイオチップ
 301 区画
 302 コントロールのスポット
 303 エLEMENTのスポット

【図1】

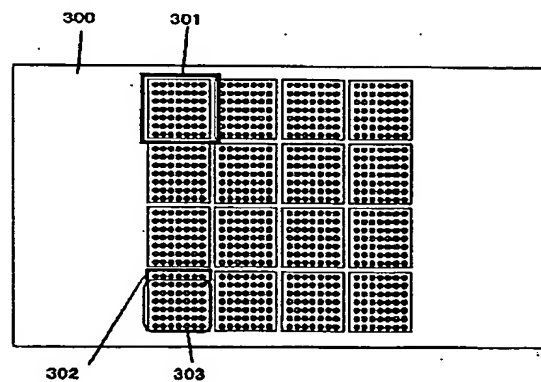


【図2】

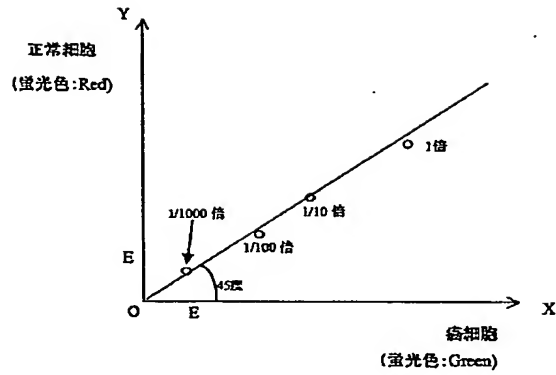
実験

遺伝子ID	正常細胞A	疾病の細胞B
1	1,234	56
2	11,224	888
3	107	3,408
⋮	⋮	⋮
⋮	⋮	⋮
n	9,753	8,905

【図3】



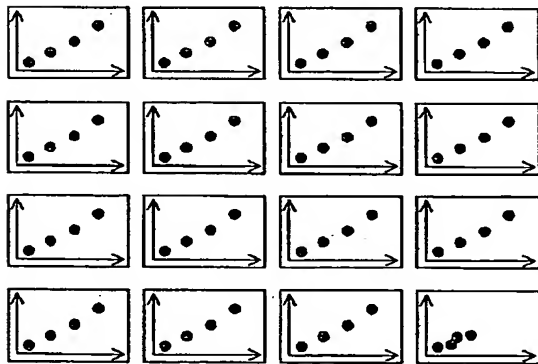
【図4】



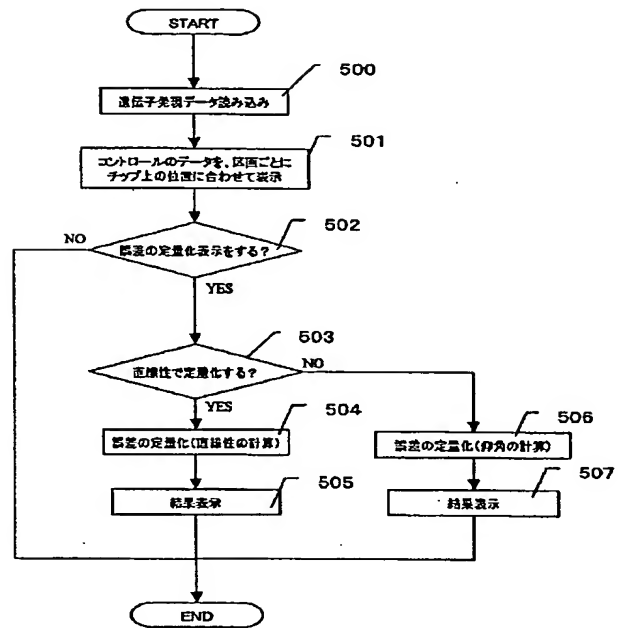
【図5】

1.1	2.1	3.1	4.1
1.2	2.2	3.2	4.2
1.3	2.3	3.3	4.3
1.4	2.4	3.4	4.4

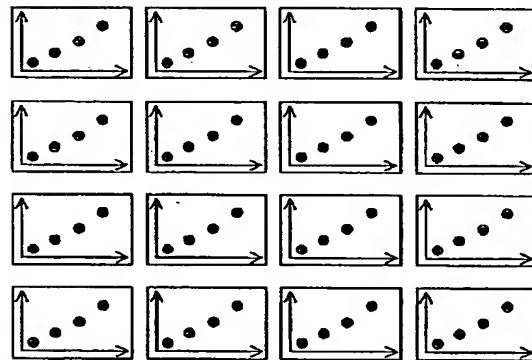
【図6】



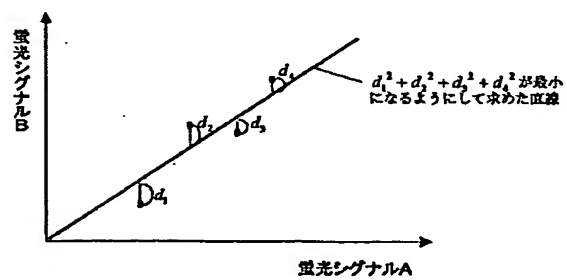
【図7】



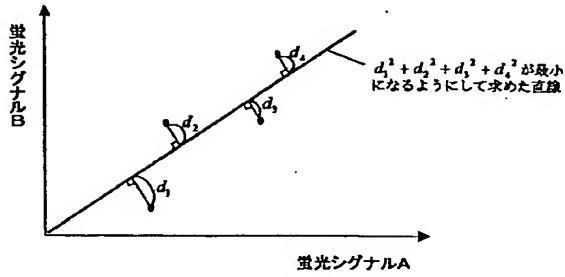
【図7】



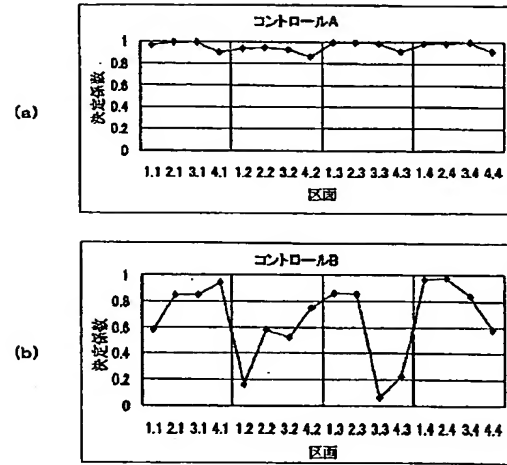
【図9】



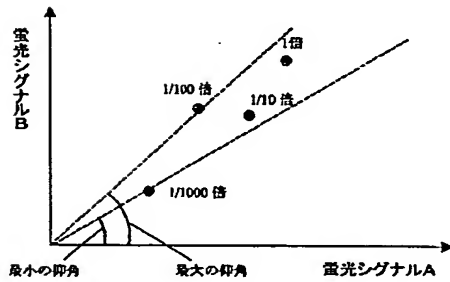
【図10】



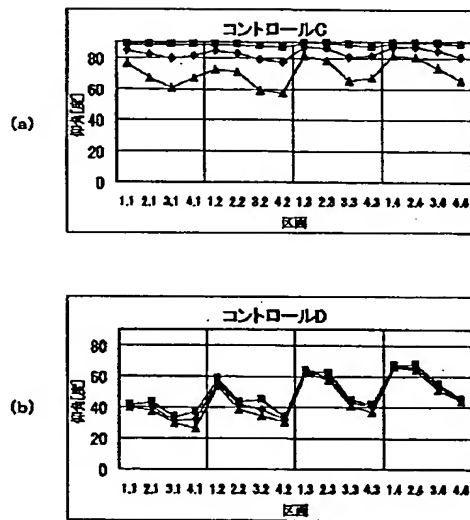
【図11】



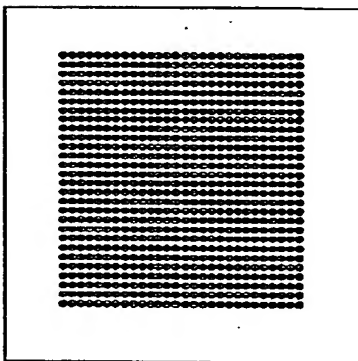
【図12】



【図13】

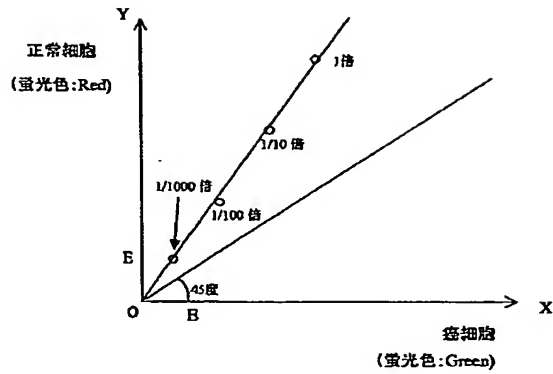


【図17】

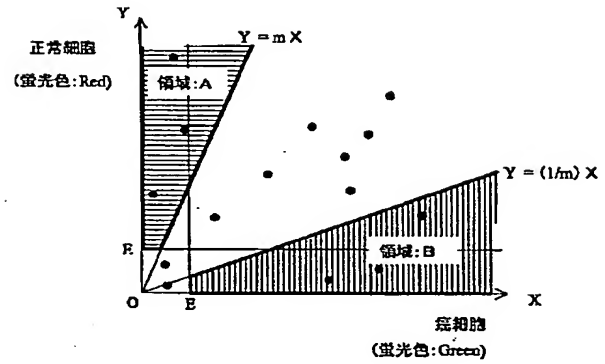


凡例: \blacktriangle : 傾角の最大 \triangle : 傾角の平均 \circ : 傾角の最小

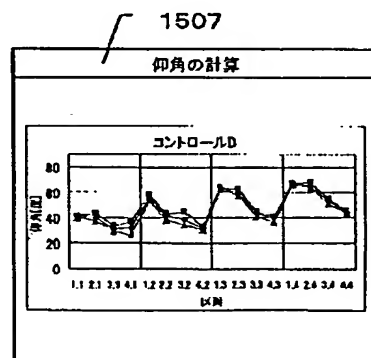
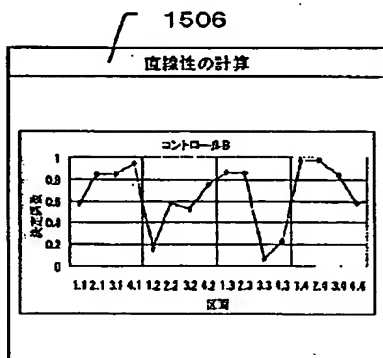
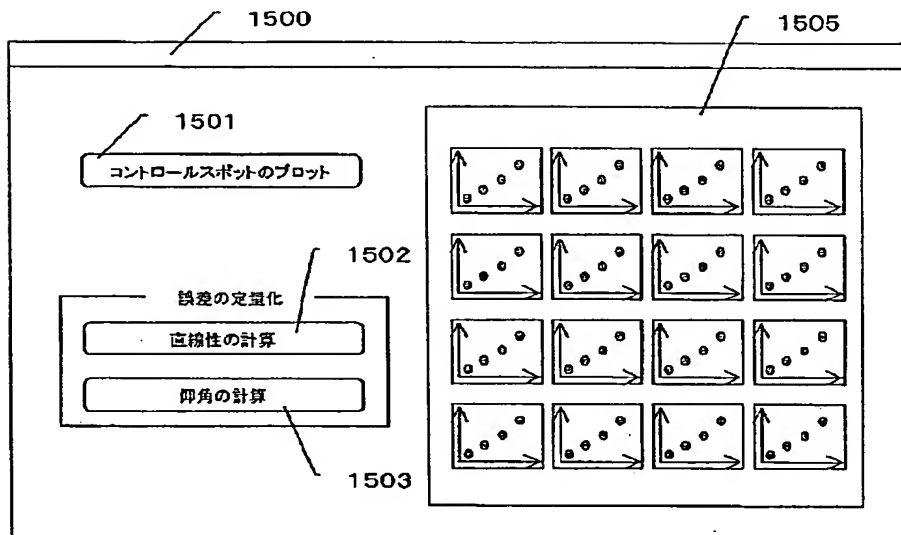
【図14】



【図16】



【図15】



フロントページの続き

(51)Int.Cl. ⁷	識別記号	F I	テーム(参考)
G 0 1 N 33/566		G 0 1 N 37/00	1 0 2
37/00	1 0 2	C 1 2 N 15/00	F
(72)発明者 野崎 康行		(72)発明者 田村 卓郎	
神奈川県横浜市中区尾上町 6 丁目81番地		神奈川県横浜市中区尾上町 6 丁目81番地	
日立ソフトウェアエンジニアリング株式会		日立ソフトウェアエンジニアリング株式会	
社内		社内	
(72)発明者 中重 亮		F ターム(参考)	2G043 BA16 CA03 DA02 EA01 KA02
神奈川県横浜市中区尾上町 6 丁目81番地			NA01 NA02 NA06
日立ソフトウェアエンジニアリング株式会			2G054 CA22 EA03 GA04 JA01 JA04
社内			4B024 AA11 AA20 CA09 HA12
			4B063 QA12 QA18 QA20 QQ42 QR32
			QR56 QR84 QS34 QS39 QX02
			QX10